

Ovládání počítače hlasem a konverzace s umělou inteligencí patří mezi klasické atributy sci-fi románů a filmů z daleké budoucnosti. Budoucnost ale nastává již dnes, hlasové ovládání počítače se stává realitou.

Radek Kubeš, radek.kubes@chip.cz

A teď budeš poslouchat!

V tomto článku najdete

Nejlepší programy pro ovládání počítače hlasem

Syntezátory lidského hlasu

Úspěšné české projekty rozpoznávání hlasu

Ovládání počítače hlasovými příkazy a syntéza řeči již několik let nepatří jen na kapitánský můstek vesmírných lodí ze Star Treku, vyzkoušet si je můžete i doma na svém počítači. Ostatně, co jiného než systém rozpoznání řeči pohání dnes zcela běžnou funkci hlasového vytáčení kontaktů

v mobilních telefonech. Přestože existují výborné vývojové produkty rozpoznání hlasu v češtině (viz rámeček), naše mateřština je pro svoji složitost a, oproti hlavním světovým jazykům, i mizivou rozšířenost v porovnání se systémy pracujícími s angličtinou značně znevýhodněna.



Najdete na Chip DVD

Typole 2.0 ■ Responding Heads 3.5 ■ VoiceMate Professional 2.15 ■ Reader 1.05 ■ Speak Text 3.0.17 ■ CIT WORDreader 3 ■ CIT MAILreader 3 ■ Opera 9.26 ■ Microsoft Voice Command 1.5 ■ Ukázková videa

Hlasové ovládání podle Microsoftu: S Vistou jen anglicky

Jedním z velkých lákadel Windows Vista je i hlasové ovládání počítače integrované přímo v operačním systému. Pro českého uživatele je hlasové ovládání použitelné jen omezeně – nepodporuje totiž češtinu, a proto není dostupné v českých verzích Visty. Není ani možné aktivovat anglickou verzi hlasového ovládání pod Vistou v češtině. Jistým řešením je pořízení anglické verze Vista Ultimate a instalace českého jazykového balíčku. Na ovládání Windows hlasem připraví uživatele průvodce, v rámci kterého se počítač naučí rozpoznávat váš hlas a přízvuk. Učení počítače probíhá i při běžném používání hlasového ovládání, kdy pomocí příkazů opravujete počítač tak, aby se naučil lépe rozpoznávat vaše příkazy. Možné je také současné ovládání počítače myší i hlasem, kterým v tomto případě nahrazujete vstup z klávesnice, jako třeba zadávání webových adres atd. Funkce hlasového ovládání Visty je velmi propracovaná, je jen škoda, že podpora češtiny zatím není na obzoru.



NAUČTE VISTU POSLOUCHAT: Tréninkový kurz naučí Windows Vista rozpoznávat hlas uživatele, zatím bohužel pouze v angličtině.

Zkušenosti s tím, jak se hlasové ovládání Visty Microsoftu povedlo, se různí. Hlasem můžete totiž nejen spouštět a ovládat běžné aplikace, ovládat myš či přepínat běžící aplikace, ale například i diktovat e-mail či dopisy ve Wordu. Humorná videa z prezentací, kde Vista neposlouchá uživatele tak docela na slovo, najdete třeba na YouTube (zadejte třeba klíčová slova „vista voice recognition“).

Windows Vista pochopitelně není prvním operačním systémem, který má vestavěné hlasové ovládání. Už v roce 2001 představila společnost IBM operační systém OS/2 Warp 4.5, který bylo možné ovládat hlasem a kterému bylo možné diktovat texty. Snaha nabídnout ovládání počítače hlasem i běžným uživatelům přímo v operačním systému je ale jistě chvályhodná a má svoji budoucnost.

univerzité v Liberci. Komerčně distribuovaná aplikace umožňuje ovládat prakticky všechny funkce počítače výhradně hlasovými povely. Standardně jsou v programu připravené povely pro spouštění a ovládání běžných aplikací, zejména MS Office, přehrávače hudby a videa, Poznámkového bloku, internetového prohlížeče a některých dalších programů. Další sadou hlasových příkazů je možné ovládat myš a klávesnici. MyVoice umožňuje rovněž diktování textu do textového editoru, a to jak po jednotlivých písmenech, tak i po celých slovech nebo frázích. Vestavěný slovník obsahuje 10 000 nejběžnějších českých slov, další můžete doplnit sami. Pokud si časem nevystačíte s přednastavenými hlasovými povely, lze program snadno naučit i další příkazy. Pro diktování se pak jistě hodí možnost zadávat třeba i velmi dlouhé, avšak často se opakující fráze jediným zvoleným příkazem.

Při používání programu MyVoice je třeba alespoň rámcově pochopit třídění povelů do skupin, naučit se příkazy pro přechod mezi skupinami a také pro základní ovládání aplikací a přepínání mezi nimi. To vše, stejně jako ovládání samotných aplikací, vyžaduje dlouhodobý nácvik výslovnosti, aby vám počítač rozuměl.

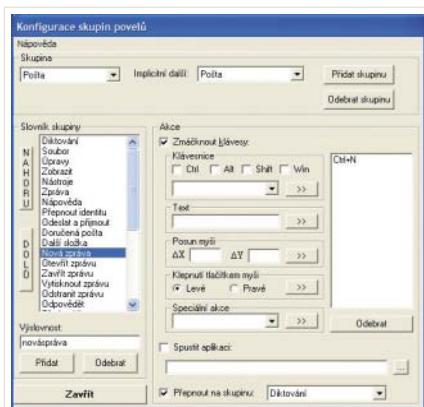
Nadstavbou programu MyVoice je aplikace MyDictate, která umožňuje diktovat celé texty, po jednotlivých slovech, do počítače. Hlasové povely slouží také pro pohyb v dokumentu. MyDictate je v základu vybaven slovníkem s více než půl milionem výrazů, další pak můžete doplňovat sami. V zájmu dosažení maximální úspěšnosti rozpoznávání hlasu je MyDictate vybaven profily mužského a ženského hlasu, navíc si pomocí namluvení skupiny výrazů můžete vytvořit profil vlastní.

Distributor programů MyVoice a MyDictate bohužel nenabízí zkušební verze programů ke stažení, nicméně na Chip DVD jsme pro vás připravili několik videoukázek z praktického používání těchto programů.

Naučte jej česky: Učící se programy

Zajímavou možností, jak naučit počítač reagovat na hlasové pokyny, je namluvení příkazů na mikrofon a definování souvisejících akcí. Aplikace poté nevychází z knihovny připravených pokynů, které se snažíte vyslovit, ale porovnává vámi nahrané vzorky s vyslovenými příkazy.

Spolehlivě funguje například aplikace **Typle**, ve které si velmi snadno, pomocí názorného průvodce, namluvíte sadu příkazů pro spouštění aplikací, nástrojů,



POVELY, KTERÉ POTŘEBUJETE: Český program MyVoice pro ovládání počítače hlasem obsahuje řadu základních povelů, další si můžete nadefinovat sami.



POSLECHNE VÁS NA SLOVO: V programu Typle si sami nahrajete hlasové příkazy a přiřadíte jim požadované funkce.

Motivací firm pro vývoj nástrojů na ovládání počítače hlasem je jistě pohodlí běžných uživatelů, avšak daleko více i umožnění plnohodnotné práce s počítačem uživatelům s různými druhy zdravotních omezení.

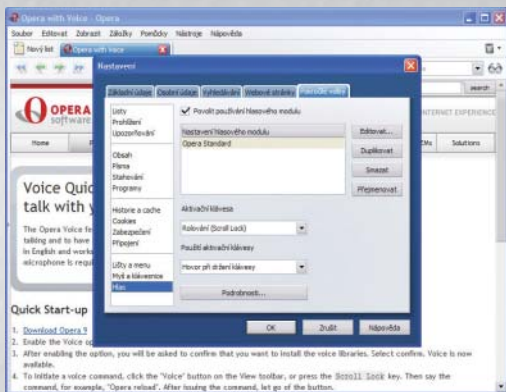
Aby vás počítač slyšel: V klidu a pomalu

K použití hlasového ovládání počítače není potřeba vlastnit žádnou speciální výbavu, stačí průměrně výkonný stroj se zvukovou kartou a běžný mikrofon, ideálně v náhlavním provedení, kdy je mikrofon blízko úst. Daleko důležitější jsou další podmínky použití rozpoznávání hlasu. Předně je třeba s počítačem pracovat v klidném a tichém prostředí. Počítač vám zatím určitě neporozumí v hlučné kanceláři, kde kromě vás mluví i vaši kolegové, vyzvánění telefony, hraje hudba nebo kam proniká hluk z ulice. Dalším předpokladem, který je ovšem možné zajistit poctivým tréninkem, je dostatečně srozumitelné vyslovování pokynů.

Mluvíte na PC česky: Posлуhač z Liberce

Prakticky jediným dostatečně funkčním projektem hlasového ovládání počítače v češtině je systém MyVoice, vyvíjený na Technické

Gesta jsou out: Ovládejte prohlížeč hlasem



HLAS PRO OPERU: Internetový prohlížeč Opera nabízí hlasové ovládání – stačí doinstalovat doplněk a umět anglicky.

Velkou novinkou ve světě internetových prohlížečů bylo před několika lety ovládání gesty, kreslenými pomocí myši. S touto funkcí přišel jako první internetový prohlížeč Opera. Inovátorský přístup si Opera udržela i v dalších verzích, kdy přišla s ovládáním internetového prohlížeče hlasovými příkazy. Stejně jako v případě Windows Vista i u Opery se netežte na možnost ovládat prohlížeč prostřednictvím povelů v češtině, nabízí se pouze angličtina.

Instalace hlasového ovládání je možná kdykoliv pomocí menu Nástroje, položky Nastavení a karty Pokročilé volby. Doplněk pro hlasové ovládání se stáhne z internetu a nainstaluje během okamžiku a nabízí

nastavení pro mužský či ženský hlas, rozdílnou intonaci, rychlost i hlasitost řeči. Hlasové ovládání aktivujete stiskem zvolené klávesy (například jinak nepotřebné klávesy Scroll Lock), vyslovíte příkaz do mikrofonu a Opera jej provede. Pomocí hlasových příkazů můžete kompletně ovládat celý prohlížeč, pohybovat se mezi stránkami, vyhledávat na internetu, pracovat s e-maily, záložkami atd. Zajímavou funkcí je čtení vybraného textu, které je ovšem pro text v češtině spíše humorné než použitelné.

Hlasové ovládání Opery je poměrně spolehlivé, úspěšnost porozumění pokynům záleží především na vaší výslovnosti angličtiny.

otevírání oblíbených internetových stránek či dokumentů, včetně případných parametrů. Tyhle poté naslouchá na pozadí systému a v případě, že zaznamená vyslovení nahrazeného pokynu, provede přiřazenou akci.

Pusťte počítač ke slovu: Slyším a odpovídám

Dalším zajímavým, bezplatně dostupným programem pro hlasové ovládání počítače je Responding Heads, využívající technologii rozpoznání řeči od Microsoftu. Aplikace zobrazí na vaší pracovní ploše mluvící hlavu, která celkem spolehlivě reaguje na vaše pokyny. Během instalace si nadefinujete typ mikrofonu, který používáte, a přečtete zkušební větu. V základu jsou v programu připraveny pouze pokyny pro ovládání samotné aplikace, další si ovšem

můžete velmi snadno nastavit sami. Responding Heads umí prostřednictvím hlasových pokynů především spouštět vybrané programy, vypsát na povel přednastavený text a přečíst označenou část textu (anglického) na webu nebo v textovém editoru. Mluvící hlavička reaguje na vaše pokyny a potvrzuje přijetí požadavku. V nastavení aplikace najdete výběr ze dvou tváří, další zmůžete vytvořit sami, pomocí editoru volně dostupného na domovském webu Responding Heads. Ani v případě Responding Heads se neobejdete bez znalosti angličtiny, váš mluvčí společník rozumí a odpovídá pouze anglicky.

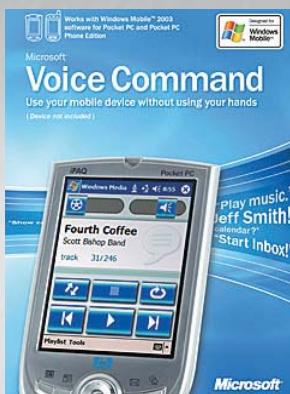
VoiceMate: Trpělivý naslouchač

Výčet programů rozpoznávajících hlasové příkazy uživatele zakončíme komerční



VIRTUÁLNÍ SPOLEČNÍK: Postavičky z programu Responding Heads reagují spolehlivě na vyslovené příkazy, stačí je to naučit.

aplikací VoiceMate. Stejně jako oba předchozí programy pracuje i VoiceMate na základě rozpoznávání pokynů a provádí jim přiřazené akce. V programu je přednastavena řada úkolů v několika kategoriích, zahrnujících například spuštění běžných aplikací, ovládání operačního systému, textového editoru, myši či klávesnice. Pomocí jednoduchého průvodce můžete do programu přidávat celé další skupiny povelů. Stačí vždy jen vybrat požadovanou akci (spuštění programu, otevření souboru atd.) a přiřadit jí hlasový příkaz napsaný tak, jak se vyslovuje. Zkušenější uživatelé jistě ocení editor skriptů, v němž si mohou naprogramovat i složitější akce s různými parametry, vykonávané počítačem na základě hlasových povelů. Hlasové příkazy je třeba zadávat v angličtině, jiný jazyk VoiceMate nerozpozná. Úspěšnost rozpoznávání hlasových povelů programem VoiceMate záleží především na použitém mikrofonu a výslovnosti.



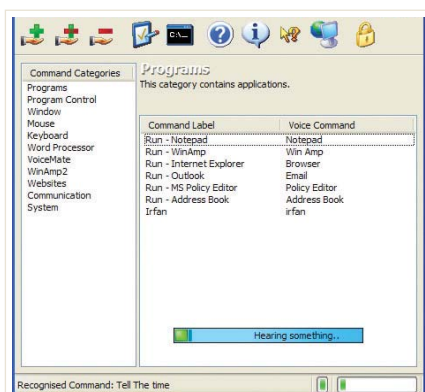
UPOVÍDANÝ TELEFON: Microsoft Voice Command přístupní v mobilních telefonech se systémem Windows Mobile hlasové ovládání i hlasový výstup.

Microsoft Voice Command: Ať vás mobil poslouchá

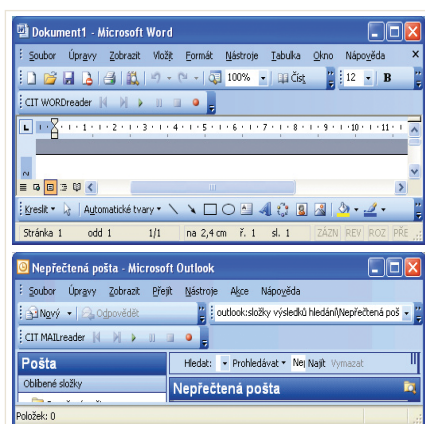
Microsoft nevyvíjí hlasové ovládání pouze pro Vistu, ale také pro stále populárnější „chytré“ mobilní telefony s operačním systémem Windows Mobile. Když pomineme hlasové vytáčení nebo vyslovení jména volajícího, které je dnes standardem i u mobilů střední třídy, nabízí Voice Command hlasové ovládání kalendáře, spuštění různých programů nebo ovládání hudebního přehrávače. Hlasové ovládání funguje také

přes Bluetooth handsfree zařízení, a to jak u sluchátek, tak například i u vestavných sad pro automobily. Aplikace zvládne rovněž syntézu hlasu, umí tedy přečíst SMS zprávu nebo hlavičku e-mailu a také například informovat uživatele o zmeškaných hovorech. Microsoft slibuje, že k použití hlasového ovládání Voice Command nebudete muset dlouhé hodiny trénovat povely. Také Voice Command funguje pouze v angličtině.

České mozky: Rozpoznávání a syntéza hlasu v ČR



ČEKÁ NA POVEL: Aplikace VoiceMate nabízí mnoho připravených akcí spouštěných hlasovými povelů, pro vytvoření dalších je připraven editor skriptů.



NECHTE SI PŘEČÍST E-MAILY: Aplikace CIT MAILreader a CIT WORDreader zvládnou přečíst e-maily i dokumenty v češtině, jejich ovládání se integruje přímo do Wordu a Outlooku.

Mluvicí počítač: Nechte si předčítat

Jednou z disciplín počítačového zpracování lidského hlasu je syntéza řeči. Tato technologie vychází vstříc opět především uživatelům s nějakou formou tělesného postižení, kteří si nemohou sami přečíst e-maily, dokumenty nebo informace na webových stránkách. Vývoj v této oblasti prošel přes používání zvukových vzorků (znáte je například z automatů na zákaznických telefonních linkách operátorů a z dalších služeb) k opravdové syntéze hlasu, který pak působí přirozeněji a srozumitelněji. Zajímavou technologickou ukázkou najdete například na adrese www.linguattec.cz/onlineservices/voice_reader/site, kde vám sličné virtuální děvče přečte zadaný text. Jedním z podporovaných jazyků je i čeština.

Z různých aplikací pro syntézu hlasu, které umí přečíst zadaný text, si můžete vyzkoušet například zdarma dostupné programy Reader nebo Speak Text. Reader funguje velmi jednoduše, stačí vložit

text do okna programu a kliknout na tlačítko přehrát. K dispozici nejsou žádná další nastavení ani funkce. Speak Reader umístí na vaši pracovní plochu postavičku kouzelníka Merlina, který, kromě toho, že netrpělivě poletuje z jednoho konce obrazovky na druhý, předčítá text vložený do schránky. Program Speak Text zvládne ještě také přečíst text uložený v TXT souboru, kouzelník rovněž umí vyslovit aktuální datum a čas. I po letmé zkoušce poznáte, že oba programy jsou spíše zajímavou než v našich podmínkách použitelnou hračkou – poradí si totiž pouze s angličtinou.

Daleko lépe zvládají češtinu komerční programy CIT MAILreader a CIT WORDreader, které slouží ke čtení dokumentů z Wordu (ve verzích 2000/XP/2003) a e-mailů ze stejných verzí Outlooku. Základní instalaci aplikace, určenou pro čtení angličtiny, je třeba rozšířit instalací doplňků pro čtení češtiny. Do Wordu i Outlooku přibudou nové nástrojové panely, které slouží jak ke spouštění čtení, tak i k obsluze funkce nahrávání počítačového čtení textu do zvukového souboru.

Budoucnost již dnes: Počítače slyší i mluví

Na příkladu několika jednoduchých, většinou i zdarma dostupných programů jsme si ukázali, že ovládání počítače hlasem, stejně jako hlasový výstup z počítače, již není hudbou budoucnosti, ale naopak realitou přítomnosti. Vývoj technologií rozpoznávání a syntézy hlasu se dostává do fáze, kdy jsou jeho výsledky použitelné a dostupné běžnému uživateli se standardním počítačovým vybavením. V případě složitých a celosvětově jen mizivě rozšířených jazyků, jako je i naše čeština, se nelze spoléhat na softwarové velmoci, ale je třeba pracovat na vlastních, pro češtinu optimalizovaných systémech.

Velkou revoluci v ovládání počítače však čekat nelze. Než se počítač bude chovat jako ve Star Treku, na to si chvíli počkáme. Ani alternativní způsoby ovládání nejsou příliš použitelné. Například se uvažovalo o ovládání počítače pomocí kamery: kamera vás pozoruje a podle toho, co děláte, vykonává příkazy. Tento systém však příliš dobře nefunguje. Máte-li webovou kamerku, můžete si zkusit zahrát aspoň on-line hru (www.zabava-online.cz/online-hry-/webkamerove-hry.html), při které k ovládání používáte právě webkameru. I když jsou hry zábavné, asi brzy zjistíte, že takto byste svůj počítač rozhodně ovládat nechtěli.

Radek Kubeš ■

■ Pro nadnárodní softwarové giganty jako Microsoft či IBM není při vývoji systémů pro rozpoznávání řeči a syntézu hlasu prioritou podpora češtiny. Má to samozřejmě svoje praktické důvody – čeština nepatří ani mezi jednoduše zpracovatelné, a už vůbec ne mezi světové jazyky. Nelze se ani spoléhat na převod některého ze systémů vyvinutých pro jiný jazyk do češtiny. Český jazyk totiž, například v porovnání s angličtinou, obsahuje několikanásobně více tvarů slov. Systémy rozpoznávání řeči v češtině proto musí pracovat s podstatně rozsáhlejšími slovníky.

To však vůbec neznamená, že by vývoj rozpoznávání a syntézy hlasu pro češtinu nebyl na vysoké úrovni. Například již od 90. let se výzkumem řečové komunikace s počítačem zabývá Fakulta mechatroniky Technické univerzity v Liberci. Cílem výzkumných projektů je pochopitelně především usnadnění ovládání počítače pohybově, zrakově či jinak handicapovaným uživatelům.



NAUČTE VISTU POSLUCHAT:

Jedním z projektů Fakulty mechatroniky TU v Liberci je i systém pro automatický přepis televizních a rozhlasových pořadů, který se v dnešní době řeší nejčastěji ručně.

Mezi projekty fakulty najdeme především aplikaci MyVoice pro hlasové ovládání počítače, o které se zmiňujeme v textu článku. Průměrně 75% úspěšnost zaznamenává systém pro automatický přepis televizních a rozhlasových pořadů, který pracuje se slovníkem přibližně 300 000 nejfrekventovanějších českých slov. Systém rozpozná osobu moderátora a poradí si i s příspěvky nahranými v rušnějším prostředí.

Českým vynálezem je zařízení I4Control, umožňující ovládání počítače pohyby očí. Speciální kamera umístěná na brýlích snímá pohyby oka a přenáší je do pohybů myši po pracovní ploše. Klikání tlačítka myši nahrazuje u I4Control mrkání. Pomocí handicapovaným uživatelům počítačů se snaží i společnost Brailcom, zaměřená na zrakově postižené.